

Accepted Manuscript

Title: Analysis of group performance with categorical data when agents are heterogeneous: The evaluation of scholastic performance in the OECD through PISA

Author: Carmen Herrero Ildefonso Mendez Antonio Villar



PII: S0272-7757(14)00019-3
DOI: <http://dx.doi.org/doi:10.1016/j.econedurev.2014.02.001>
Reference: ECOEDU 1474

To appear in: *Economics of Education Review*

Received date: 24-4-2013
Revised date: 5-2-2014
Accepted date: 6-2-2014

Please cite this article as: Carmen Herrero, Ildefonso Mendez, Antonio Villar, Analysis of group performance with categorical data when agents are heterogeneous: The evaluation of scholastic performance in the OECD through PISA, *Economics of Education Review* (2014), <http://dx.doi.org/10.1016/j.econedurev.2014.02.001>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Analysis of group performance with categorical data when agents are heterogeneous: The case of compulsory education in the OECD

Carmen Herrero
University of Alicante, Ivie & IAE

Ildefonso Mendez*
University of Murcia

Antonio Villar
Pablo de Olavide University & Ivie

Abstract

This paper analyses the evaluation of the relative performance of a set of groups when their outcomes are defined in terms of categorical data and the groups' members are heterogeneous. This type of problem has been dealt with in Herrero and Villar (2013) for the case of a homogenous population. Here we expand their model controlling for heterogeneity by means of inverse probability weighting techniques. We apply this extended model to the analysis of compulsory education in the OECD countries, using the data in the PISA. We evaluate the relative performance of the different countries out of the distribution of the students' achievements across the different levels of competence, controlling by the students' characteristics (explanatory variables regarding schooling and family environment). We find that differences in mathematical and reading abilities across OECD countries would lower by between 40% and 50% if their endowment of students' characteristics would be that for the OCDE average.

Keywords: Group performance, compulsory education, heterogeneity, categorical data, inverse probability weighting.

JEL codes: I24, C14.

*Corresponding author. E-mail: ildefonso.mendez@um.es. Mailing address: Departamento de Economía Aplicada, Facultad de Economía y Empresa, 30100, Espinardo (Murcia), Spain. Phone: +34 868883732. Fax: +34 868883745.

We evaluate relative performance with heterogeneous groups and categorical outcomes.

We apply our model and evaluate OECD schoolchildren's mathematical and reading abilities.

Students' characteristics account for one half of the country differences in PISA.

We show how the relative position of countries changes as we expand the covariates.

We analyze how relative performance has changed in the first decade of the 21st Century.

Accepted Manuscript

Analysis of group performance with categorical data when agents are heterogeneous: The case of compulsory education in the OECD

Abstract

This paper analyses the evaluation of the relative performance of a set of groups when their outcomes are defined in terms of categorical data and the groups' members are heterogeneous. This type of problem has been dealt with in Herrero and Villar (2013) for the case of a homogenous population. Here we expand their model controlling for heterogeneity by means of inverse probability weighting techniques. We apply this extended model to the analysis of compulsory education in the OECD countries, using the data in the PISA. We evaluate the relative performance of the different countries out of the distribution of the students' achievements across the different levels of competence, controlling by the students' characteristics (explanatory variables regarding schooling and family environment). We find that differences in mathematical and reading abilities across OECD countries would lower by between 40% and 50% if their endowment of students' characteristics would be that for the OCDE average.

Keywords: Group performance, compulsory education, heterogeneity, categorical data, inverse probability weighting.

JEL codes: I24, C14.

1 Introduction

We consider here an evaluation problem in which we have to compare the relative performance of several groups, out of the distribution of the achievements of their members in a set of ordered categories. Think for instance of the comparison of the health situation of different countries out of the distribution of the population in four or five health statuses (e.g. from "excellent" to "very bad"). The key elements of the problem are, therefore, the presence of several groups, the qualitative nature of the outcome variable, which resolves into a given set of ordered categories, and the focus on relative performance.

This type of problem has been addressed recently by Herrero and Villar (2013). They start by considering pairwise comparisons between groups in terms of the probability that an agent picked at random in one group belongs to a higher category than an agent randomly chosen in the other. Then they extend those comparisons to all groups involved, by taking into account both direct and indirect relations. As a result they obtain an evaluation function that corresponds to the dominant eigenvector of a matrix that describes all those comparisons (see below). This evaluation function is characterized in terms of some ethical and operational properties.

An implicit assumption in their model is that groups are homogeneous so that the distribution of the outcome variable is the sole relevant information. Yet one might be interested in evaluating not only the observed outcomes but also the extent to which those outcomes reflect diverse structural characteristics of the population that affect the agents' performance. This may well be the case in the example mentioned above (comparing the health situation of different countries), regarding the influence of aspects such as age or wealth in the final outcomes. To deal with this type of evaluation we need a methodology that permits making comparisons in terms of a common set of characteristics. This is the key point of this paper. More specifically, we combine here the original model in Herrero and Villar (2013) with inverse probability weighting (IPW) techniques that permits one controlling for differences in the distribution of the determinants of the outcome variable.

Using this methodological approach we obtain a covariate-adjusted evaluation that allows isolating the impact of the selected explanatory variables, by comparing this evaluation with the unadjusted one. In that way we can separate the part of the observed differences that is explained by the covariates and the part which cannot be accounted for. The covariate-adjusted eigenvector tells us about the relative performance of the groups once their conditioning variables have been equalized. Comparing the covariate-adjusted and the unadjusted evaluations permits one to estimate the impact of the latent variables on the relative performance.

The interpretation of the differences between both evaluations depends on the problem at hand and, in particular, on the choice of the explanatory variables. In this respect our analysis is reminiscent of the "equality of opportunity" literature, as covariate-adjusted values might be interpreted as an expression of the differential "effort", whereas the unadjusted values would reflect the in-

terplay of both effort and "opportunity".¹ Yet, this model does not provide a "measure" of equality of opportunity, as our comparison deals with relative performance both in the adjusted and unadjusted evaluations.

We apply this extended model to the evaluation of compulsory education in the OECD using the data provided by the Program for International Students Assessment (PISA). We evaluate the performance of schoolchildren regarding mathematical ability, out of the 2012 data set (the last one available). Our evaluation involves the estimation of the impact of the students' environment (parental and school characteristics) on the final scores. Comparing the adjusted and unadjusted evaluations allows concluding that the set of explanatory variables accounts for almost 50% of the differences in the relative performance. We also consider how those differences have evolved in the first decade of the 21st Century, by comparing the results in 2012 and those in 2003. We find that differences in students' mathematical ability across OECD countries have decreased during that period, particularly so for European OECD countries.²

We also analyze the evolution of country differences in reading competence using the 2009 and 2000 PISA reports, that focused on that subject. We find that the differences in students' reading ability across OECD countries lowered during the first decade of the 21st Century to a greater extent than the differences in mathematical competence did. The countries that improve (lower) their relative position in one of the two competences are likely to be those that improve (lower) their relative position in the other competence. Finally, we find that students' environmental factors also account for between 40% and 50% of country differences in relative performance in reading ability.

The paper is organised as follows. Section 2 presents the formal model whereas Section 3 applies it to the results on mathematical competence out of the data in the PISA (2003 and 2012). Section 4 gathers a few final comments.

2 The Model

Consider a set of g groups or societies, $G = \{1, 2, \dots, g\}$, each of which consists of n_i agents, $i \in G$. We want to compare the relative performance of those groups with respect to a given aspect, when their achievements are given in

¹Equality of Opportunity (EOp) is one of the most prominent concepts of distributive justice. The key idea behind this concept is that the concern about inequality should not focus on the equality of outcomes but rather on the existence of a common *playing field* for all people. From this perspective agents' outcomes can be regarded as deriving from two different sources: *effort and opportunity*. Effort refers to people's decisions whereas opportunity refers to the agents' external circumstances. A fair society is one in which final outcomes do not depend much on the agents' external circumstances, that is, a society in which all people share similar opportunities. In that society outcome differences are basically determined by the agents' preferences and effort and not by aspects that are beyond their control and responsibility (see Arneson, 1989; Cohen, 1989; Roemer, 1993, 1998; Fleurbaey, 2008).

²Here again our analysis is very close to that of equality of opportunity in education. See on this respect Peragine and Serlenga (2008), Lefranc et al. (2008), Chechi and Peragine (2010), Calo-Blanco and Villar (2010), OECD (2010 b), Villar (2012), Calo-Blanco and García-Pérez (2013).

categorical terms. More precisely, we assume that there is a set of categorical positions, $H = \{h_1, \dots, h_s\}$, ordered from best to worst, $h_1 \succ \dots \succ h_s$ (health statuses, educational levels, age intervals, professional positions, etc.). Each group presents a given distribution of achievements across those categories. Our goal is comparing their relative performance, taking into account the role of the differences in the structural characteristics that may influence the outcome variable. To do so we divide the evaluation problem into two parts. First, we assume that all groups are homogeneous regarding those characteristics, so that the evaluation only takes into account their relative achievements. The key point here is how to make systematic comparisons out of qualitative data. Second, we consider that groups are heterogeneous and provide a method to control for such heterogeneity.

2.1 The evaluation formula when groups are homogeneous

Let $G = \{1, 2, \dots, g\}$ stand for a set of g groups under the assumption that they are homogeneous with respect to the aspect under evaluation. Let a_{ir} , for $i = 1, \dots, g$, $r = 1, \dots, s$, be the proportion of people of group i in position r , and let A stand for the matrix that collects all those values (that is, the i th row of matrix A describes the distribution of achievements of group i across the different categories, in terms of relative frequencies).

An evaluation problem, or simply a problem, can be summarized by that matrix A , under the assumption that the set of groups, G , and the set of categorical positions, H , are given, and that we focus on the relative frequencies of the agents across the categorical positions, independently on the size of the groups (a property known as *group replication invariance*). Our target is to define a suitable evaluation function that enables comparing the relative performance of the different groups.

Given an evaluation problem A we say that group i dominates group j when it is more likely that an agent chosen at random from i occupies a higher position than an agent chosen at random from j . Let p_{ij} be the probability that an agent from group i occupies a higher position than an agent from group j , and let $e_{ij} = 1 - p_{ij} - p_{ji}$ (that is, the probability that an individual in i , chosen at random, belongs to the same category that an individual in j , randomly chosen). Note that, those probabilities can be easily computed as follows:

$$\begin{aligned} p_{ij} &= a_{i1}(a_{j2} + \dots + a_{js}) + a_{i2}(a_{j3} + \dots + a_{js}) + \dots + a_{i,s-1}a_{js} \\ e_{ij} &= a_{i1}a_{j1} + \dots + a_{is}a_{js} \end{aligned}$$

In pairwise comparisons, the quotient p_{ij}/p_{ji} tells us the relative advantage of group i with respect to group j . That is, $p_{ij}/p_{ji} > 1$ implies that people in i have advantage over people in j , and viceversa.³

³Liebertson (1976) in a similar vein introduces the Index of Net Difference, $ND(i, j) = |p_{ij} - p_{ji}|$ to inform about inequalities between two groups. If $ND(ij) = 0$, then $p_{ij} = p_{ji}$ that is, it is equally likely, given an individual chosen at random in any of the two groups that the individual in i is at a better position than the individual in j than the other way around.

Remark 1 When the distribution of the population in group i stochastically dominates that of group j , we have that $p_{ij} > p_{ji}$.⁴

Note that, when there are more than two groups, pairwise comparisons only cover part of the relevant domination relationships. This is so because in that case one has to take into account not only direct dominance relations but also the indirect ones. That is, the relative position of group i with respect to group j also depends on how those groups relate to third parties.

Herrero and Villar (2013) introduce a summary measure of relative achievements that can take into account all those relations. They define the *relative advantage* of group i with respect to group j , RA_{ij} , as the ratio between the probability that i dominates j and the sum of the probabilities that group i be dominated by other groups (a number between 0 and $(g - 1)$). Formally:

$$RA_{ij} = \frac{p_{ij}}{\sum_{k \neq i} p_{ki}}$$

That is, RA_{ij} is directly proportional to the probability of i dominating j and inversely proportional to the sum of the probabilities of i being dominated by some other group. Observe that when there are only two groups, $RA_{ij} = p_{ij}/p_{ji}$, whereas this value changes in the presence of more groups. A summary measure of the overall advantage of a given group can thus be obtained as a weighted average of the relative advantages with respect to all groups. That is,

$$RA_i = \sum_{j \neq i} \lambda_j RA_{ij}$$

where λ_j is the weight attached to group j .

It is natural to look for a weighting system consistent with the evaluation of relative advantages. That is, a weighting system such that $\lambda_j = RA_j$. That means that one has to find a vector $\mathbf{v} = (v_1, v_2, \dots, v_g)$ such that:

$$v_i = \sum_{j \neq i} v_j RA_{ij} = \sum_{j \neq i} \frac{v_j p_{ij}}{\sum_{k \neq i} p_{ki}} \quad (1)$$

Herrero and Villar (2013) show that such a vector always exists, it is strictly positive and unique (up to normalization), and has an interesting number of additional properties. Moreover, that vector \mathbf{v} is easy to compute, since it corresponds to the Perron eigenvector of the following matrix:

$$Q = \begin{pmatrix} g - 1 - \sum_{i \neq 1} p_{i1} & p_{12} & \cdots & p_{1g} \\ p_{21} & g - 1 - \sum_{i \neq 2} p_{i2} & \cdots & p_{2g} \\ \cdots & \cdots & \cdots & \cdots \\ p_{g-1,1} & p_{g-1,2} & \cdots & p_{g-1,g} \\ p_{g1} & p_{g2} & \cdots & g - 1 - \sum_{i \neq g} p_{ig} \end{pmatrix}$$

The other extreme case is when $ND(i, j) = 1$, which happens whenever all individuals in one of the groups are at better positions than those in the other group. Intermediate positions provide with values of $ND(i, j)$ between 0 and 1.

⁴Lefranc, Pistolesi and Trannoy (2008, 2009) make use of stochastic dominance for income distributions to compare equality of opportunity among two different groups of people.

The interpretation of the components of matrix Q is the following. Off-diagonal elements (elements in place ij with $i \neq j$) are simply p_{ij} , that is, the probability that an individual chosen at random in group i is at a higher position than an individual chosen at random in group j . Thus all off-diagonal elements capture the relative dominance between pairs of groups. As for the elements in the diagonal, the element jj provides the probability of someone chosen at random in group j to be at least in a position as good as (or better than) anyone in any other group.

It is easy to check that matrix Q is a Perron matrix all whose columns add up to $(g-1)$. Therefore, assuming that matrix Q is irreducible, there is a unique eigenvector, $\mathbf{v} \gg \mathbf{0}$, absorbent, and such that its components add up to g . The components of such an eigenvector satisfy equation [1], and thus, provide the evaluation we are looking for. Vector \mathbf{v} can thus be regarded as a summary measure of the relative performance of the different groups.

2.2 Controlling for heterogeneity

We now show how to combine this model with inverse probability weighting (IPW) techniques that permit one controlling for differences in the distribution of the determinants of the outcomes in the different groups. We obtain in this way a covariate-adjusted eigenvector that provides an evaluation of the relative performance once the impact of the differences in the distribution of the covariates has been cancelled. Comparing covariate-adjusted and unadjusted evaluations tells us about the influence of the explanatory variables in the observed performance.

The IPW estimators are easy to implement, allow for an undetermined amount of heterogeneity in the estimates, and make no assumption on the distribution of the outcome variable H . Additionally, it has been shown in the treatment effects literature that the IPW estimators provide consistent and in some cases asymptotically efficient estimates of the parameter of interest under fairly standard regularity conditions. Furthermore, Busso et al. (2012) showed that the IPW estimators exhibit the best overall finite sample performance among the broad class of treatment effect estimators. This is particularly relevant in the current context since estimation samples are of modest size in many empirical applications.

The goal is to ensure the same distribution of the covariates (X) in each category (h) of each group (r) used to calculate the eigenvector. We first choose one of the $C = g \cdot s$ subsamples in which the total sample is partitioned, labeled c_r , as the reference sample, i.e. that whose distribution of covariates is to be used in the remaining $C - 1$ subsamples. Alternatively, we could use the total sample for group g or the overall sample for all the groups as the reference sample. Later in this section we analyze how the reference sample affects the outcome of the evaluation tool and we also provide some insights on how to select the reference sample.

Next, we generate a set of dummy indicator variables Z_c that equal one if an observation belongs to the reference sample and zero if it belongs to sub-

sample c , for $c \in C, c \neq c_r$. We then estimate the conditional probability of being in the reference sample given X , i.e. $p_c(x) = P(Z_c = 1/X = x)$ for each observation in subsample c , for $c \in C, c \neq c_r$. This variable is known as the *propensity score* in the treatment effects literature. The research value of the propensity score rests on its power to solve the dimensionality problem, since adjusting for between-groups differences on a high dimensional vector of covariates can be either difficult or impossible. Rosenbaum and Rubin (1983) show that the propensity score captures all of the variance on the covariates relevant for adjusting between-group comparisons, that is, treated ($Z_c = 1$) and control ($Z_c = 0$) units with the same value of the propensity score have the same distribution of the elements in X .

The propensity score can be estimated by means of a simple binary choice model like a logit or a probit model or by nonparametric methods like power series regression. The distribution of the covariates in a particular subsample is changed for that in the reference sample by simply introducing the appropriate weighting function λ_c . Formally, let $g(X)$, $g(X/Z_c = 1)$ and $g(X/Z_c = 0)$ be the joint density of X in the estimation sample, in the reference subsample and in subsample c , respectively, and observe that by definition,

$$g(X) = \frac{g(X/Z_c = j) P(Z_c = j)}{P(Z_c = j/X)}, \text{ for } j = \{0, 1\}.$$

Then, it follows that:

$$\underbrace{\frac{P(Z_c = 1/X) (1 - P(Z_c = 1))}{(1 - P(Z_c = 1/X)) P(Z_c = 1)}}_{\lambda_c} \cdot g(X/Z_c = 0) = g(X/Z_c = 1), \quad (2)$$

where $P(Z_c = 1)$ is the proportion of observations from the reference sample in the estimation sample. This equation suggests a simple three-step method to change the distribution of X in each subsample for that in the reference sample. First, get an estimate of the propensity score for each observation in the sample. Second, plug the estimated propensity score and the proportion of observations from the reference sample in the estimation sample into the sample analog of λ_c to obtain an estimate of the weighting function.⁵ Next, use the estimated λ_c to weight observations of subsample c and calculate, for each group, the proportion of observations in each category. The covariate-adjusted eigenvector is obtained by simply applying the evaluation tool due to Herrero and Villar (2013) to the new percentages.

This weighting scheme works by weighting-down (-up) the distribution of 1s of the dummy indicator variable $(1 - Z_c)$ for observations in subsample c for those values of the elements of X that are (over-) under-represented among observations in the reference sample. The following overlap assumption on the joint distribution of Z and X is necessary for the estimation problem to be well defined: $0 < P(Z_c = 1/X) < 1$, for $c \in C$. This common support condition

⁵The weights are normalized so that they add up to one within each subsample.

states that for a given value of X there is some fraction of the estimation sample in the reference sample and in subsample c to be compared. Lack of overlap as well as estimated propensity scores close to one can lead to imprecise estimates of λ_c . To overcome this limitation we follow Crump et al. (2012). They propose a systematic approach to addressing lack of overlap in estimation of average treatment effects by characterizing optimal estimation subsamples for which the average treatment effect can be estimated most precisely. Remarkably, the optimal rule depends solely on the propensity score in most cases. We apply the procedure in Crump et al. (2012) for the treatment effect on the treated, the parameter whose weighting function is that in [2], to each of the comparisons that we perform.

That could be argued that only the observations of the reference sample that belong to all the optimal estimation subsamples defined according to Crump et al. (2012) should be used to calculate the covariate-adjusted eigenvector. However, that requirement can only be satisfied if C is not very large and the subsamples being compared do not differ to a great extent in the distribution of the variables in X . Otherwise, the cost in terms of sacrificed external validity would be higher than the improvement in internal validity.

The choice of the reference sample affects the relative importance of each category within each group (a_{ij}), the probability that an agent from group i occupies a higher position than an agent from group j (p_{ij}), the ratio between any two p 's and, thus, the outcome of the evaluation process (v_i).⁶ The more selected the reference sample, the higher the cost in terms of sacrificed external validity. That is the case for the lowest and highest proficiency levels, particularly so in the latter case for the less developed OECD countries. We use the overall sample of OECD countries as the reference sample in our application. This way we analyze how the relative scholastic performance of the OECD countries would change if their endowment of environmental variables would be that for the OCDE average.⁷

The covariate adjusted matrix corresponds to the homogenous groups case. That is, the information has been transformed so that it describes the outcomes that would have been obtained if all groups were similar regarding the explanatory variables. The covariate-adjusted eigenvector, \mathbf{v}^* , tells us the relative differences in the groups that are not accounted for those variables. And, consequently, the difference between the eigenvector obtained with the original data and the covariate-adjusted one tells us about the impact of leveling the environmental variables in the evaluation of the relative performance of the different groups.

This model is applied in the next Section to evaluate schoolchildren scholastic performance across 30 OECD countries. We do so by considering the distrib-

⁶Conversely, the ratio between any two λ 's is invariant to the selection of the reference sample.

⁷Our results remain qualitatively largely unchanged when we use particular subsamples as the reference sample. In general, we find that the more selected is the reference subsample, the higher is the difference that accounting for covariates makes with respect to the unadjusted eigenvector.

ution of the students in five categories of mathematical competence (the six determined by PISA, after merging levels 5 and 6). This leaves us with 150 subsamples to compare. Moreover, there exists a large heterogeneity in the distribution of the variables in X both between and within OECD countries. For these reasons, we cannot restrict the reference subsample to those observations satisfying the optimal selection rule in all the comparisons. The same holds when we restrict the analysis to the 21 OECD European countries.

3 The evaluation of compulsory education in the OECD through PISA

Compulsory education is probably the most basic instance of social insurance as it guarantees minimal levels of knowledge to all citizens in a given society, which in turn conditions their opportunities regarding access to the labour market, further education, and the extent of social interactions. Most OECD countries have established a ten- to eleven-year period of compulsory education (from 6 to 15 or 16 years). It is important to know how effective this education is, how different are the educational outcomes both between and within countries, and how much those differences depend on the countries' observable characteristics.

Here we apply the model presented above to evaluate compulsory education in the OECD countries out of the data provided by the Programme for International Student Assessment (PISA). This Programme provides the broadest dataset for the evaluation of schoolchildren performance and the characteristics of their schooling and family environments. It is a triennial worldwide test of 15-year-old schoolchildren's scholastic performance, the implementation of which is coordinated by the OECD. PISA surveys started in 2000 with the aim of evaluating the students' ability, about the end of compulsory education, in three different domains: mathematical, mathematics and science.

Every period of assessment specialises in one particular category, but it also tests the other two main areas studied. The 2012 report is the most recent one and it focused on mathematical ability, as the 2003 report did. We use data in PISA 2012 to analyze the different performance of educational systems regarding schoolchildren's mathematical ability across OECD countries.⁸ The focus of our analysis is threefold. First, we investigate the relative performance of OECD countries in 2012 comparing the distribution of mean score tests and their relative performance, as measured by the corresponding (unadjusted) eigenvector. Second, we estimate the students' relative achievements once their external circumstances have been equalized, using two different sets of explanatory variables. And third, we provide a rapid overview of the changes experienced in 2012 with respect to 2003. Given this purpose, we consider 30

⁸One of the assets of the PISA report is that it provides a unified scoring system to evaluate the performance of 15-year-old students in very different countries. The units of those scores are set with respect to the values obtained in the 2000 wave of the report, by taking a value of 500 for the average of the OECD Member States with a standard deviation of 100.

out of the 34 OECD countries, since Chile, Estonia, Israel and Slovenia did not participate in the 2003 wave.

We also present evidence on the evolution of country differences in reading competence using the 2009 and 2000 PISA reports, that focused on that subject. When doing so, we restrict the analysis to the 27 OECD countries that participated in the four PISA reports and that provide information on the variables used in the analysis.⁹

3.1 Comparing the relative performance of OECD educational systems in 2012

The PISA report provides a classification of the students in six different levels of mathematical competence, from Level 1 (the lowest level) to Level 6 (the highest one). Those levels are defined in terms of the capacity of the students to master certain cognitive processes and operationalized in terms of ranges of the scores obtained by the students (see Figure 1.2.12 in OECD (2010a) for details). Table 1 summarizes the scoring intervals that parameterize those levels of competence.

Insert Table 1 here

We combine levels 5 and 6 of mathematical competence into a unique upper level since the share of students in the sixth level is quite low in most countries and it does not allow us to obtain accurate estimates of the propensity score for the students in that level.¹⁰ The reduction in the number of levels of mathematical competence leaves almost unchanged the unadjusted eigenvectors but it substantially increases the stability of the covariate-adjusted eigenvectors as we increase the number of covariates used to estimate the propensity score. The distribution of students within the five levels of mathematical competence in 2012 is presented in Table 2.

Insert Table 2 here

In order to compare the performance of educational systems out of the distribution of outcomes in Table 2 we rely on the model of Section 2. Now *groups* are OECD countries, *members* are the 15-year old students within each country, and *categories* correspond to the five levels of mathematical competence. Table 3 below presents the eigenvector for the year 2012 and compares those values with the mean scores of the PISA tests.¹¹ We present both unadjusted

⁹Japan, the Slovak Republic and Turkey are excluded from the estimation sample in this comparative exercise.

¹⁰Less than 1% of the students have a reading competence of level 6 in 26 out of the 29 OECD countries.

¹¹Test results are not presented in PISA as point estimates. Rather, PISA reports student performance through five plausible values that can be defined as random values from the posterior distribution of an student's performance (see OECD (2009) for details). As indicated in OECD (2009), we perform our statistical analysis independently on each of the five plausible values. The final eigenvector is the average of the eigenvectors obtained for each of the five plausible values. Anyway, the resulting eigenvectors are almost identical to

and covariate-adjusted eigenvectors (indeed we present two different covariate adjusted vectors, to be explained below). The eigenvectors have been normalized so that the sum of its components equals the number of components. That is, values above one indicate that the country performs over the mean, while components below one are to be interpreted as performing below the mean. Equivalently, the mean scores of the PISA tests have been normalized by setting its average equal to one.

Insert Table 3 here

The unadjusted eigenvectors and the mean scores of the PISA tests yield similar results in terms of ranking but rather different pictures in terms of cross-country differences, with the distribution of mean score values being far more concentrated around the mean than that of eigenvector components. Indeed, while the Kendall's rank order correlation coefficient between the two evaluation measures is 0.93, the coefficient of variation of the eigenvector components is 6.7 times that of the mean scores.

The normalized mean score underestimates the relative advantage of the countries that perform best at mathematics according to the eigenvector and overestimates the relative position of the countries that perform worst. For example, the normalized mean score of Korea, Japan, Belgium and Germany, the four countries at the top of the distribution according to the eigenvector, are 46%, 35% and 25% lower than the corresponding eigenvector components, respectively. Conversely, the mean scores of Sweden, Turkey, Greece and Mexico, the four countries at the bottom of the distribution according to the eigenvector, are 37%, 103.2%, 106% and 307% higher than the corresponding eigenvector components, respectively. The difference between the two evaluation methods is of modest size for the countries that perform close to the mean.

3.2 Accounting for differences in characteristics

We now move to the covariate-adjusted eigenvectors in column (1). These were obtained by using the set of students' external factors that is common to both the 2003 and the 2012 reports in the estimation of the propensity score.¹² In this specification we control for children's sex and for the employment status and the educational level of their parents by means of a set of dummy indicator variables that inform on whether the father or the mother are employed and whether their highest educational level is the secondary or the tertiary level as defined in the International Standard Classification of Education (ISCED). We also control for whether at least one of the parents was born in a country different from their current country of residence and for whether there are at least one hundred books at home or not.

Regarding their schooling environment, we control for the share of the school total funding for a typical school year that comes from the government and

those obtained using the average of the five plausible values as the summary measure of the students' performance.

¹²The propensity scores are estimated using binary logit models.

whether the assessments of students are used to make judgments about teachers' effectiveness or not.¹³ All these familiar and schooling factors are relevant determinants of international differences in students' educational achievement according to Hanushek and Woessmann (2011). Tables 4 and 5 summarize the distribution of the familiar and schooling determinants for the 30 OECD countries, respectively.

Insert Table 4 here

Our estimates attest that accounting for heterogeneity makes a relevant difference. On the one hand, the coefficient of variation of the eigenvector lowers by 49%, once we control for differences in characteristics. That is, the explanatory variables in (1) account for almost half of the variation in schoolchildren's scholastic relative performance in mathematical ability across OECD countries. On the other hand, the Kendall's correlation coefficient between the unadjusted and the covariate-adjusted eigenvectors is of only 0.37. This indicates that differences in characteristics across OECD countries account for approximately two thirds of the sorting of OECD countries that results from the unadjusted eigenvector.

The relative advantage of children coming from the countries that do best at mathematics according to the unadjusted eigenvector lowers once we control for heterogeneity. That is the case for Korea, Japan, Belgium, Germany and Netherlands, whose relative advantage lowers between 20% and 50%. Conversely, the relative position of the countries that do worst according to the unadjusted eigenvector improves once we control for X . Portugal, Turkey, Greece and Mexico improve their relative performance by approximately 55%, 91%, 63% and 212%, respectively. Poland stands out for being the only country that performs over the mean according to the unadjusted eigenvector that improves its relative position once we control for X . Germany, Finland, New Zealand and Canada move from performing over the mean to performing below the mean once we control for students' external circumstances. The opposite holds for the Slovak Republic, Portugal and the United States.¹⁴

One of our main findings is that almost half of the inequality in students' mathematical ability across OECD countries is explained by country differences in students' family and schooling characteristics. How much of the remaining differences can be accounted for by introducing additional explanatory variables? To investigate this issue we have included in Table 3 another covariate-adjusted eigenvector in column (2), estimated by expanding the set of controls in specification (1) with some determinants of students' achievements.¹⁵ This allows us to analyze how the coefficient of variation of the eigenvector components

¹³The relevance of government funding is preferred to the indicator of whether the school is public or private because of the large number of missing values in the latter variable in the 2003 report.

¹⁴The relative position of Spain slightly improves once we control for X . This result is in line with that in Ciccone and Garcia-Fontes (2009). They found that there is an increase in Spain's PISA scores relative to the rest of Europe when parental schooling is accounted for.

¹⁵Those are present only in the 2012 report and cannot be considered when comparing the results in 2012 and in 2003.

changes as we expand the set of covariates. In particular, in (2) we additionally control for whether the school is located in a large city, i.e. a city with over one million people, or not, for the percentage of teachers fully certified by the appropriate authority in the school, whether students are grouped by ability or not, whether the school monitors the practice of teachers or not and whether the principal, the department head or the teachers have the main responsibility for hiring teachers or not. Table 5 summarizes the distribution of these schooling characteristics across the 30 OECD countries. As before, these variables are relevant determinants of students' achievement tests according to Hanushek and Woessmann (2011).

Insert Table 5 here

The difference between the two covariate-adjusted eigenvectors is of modest size, below 4% for 22 out of the 30 components. Moreover, the coefficient of variation remains largely unchanged when we expand the set of controls. This finding suggests that there is a part of the difference in the evaluation of students' mathematical ability across OECD countries that cannot be explained by the differences in the characteristics that the theory suggests. That is to say, there is a set of unobservable factors that induce the existence of international differences, beyond the action of the standard determinants. Those unobservable variables may refer to cultural and organizational factors, such as the design of the educational system, the commitment of families and society with education, the implementation of educational policies, the teachers' attitudes and involvement, and, of course, the average effort of the students. The low sensitivity of the coefficient of variation of the covariate-adjusted eigenvector, with respect to the inclusion of five relevant additional covariates, points out the relevance of those factors. This is supported by the fact that the same type of results appears between the regions of some countries.¹⁶

3.3 The evolution of relative performance

Let us now comment on how those values have evolved between 2003 and 2012, so that we can have an idea of the dynamics of the educational systems. To ease the discussion of the results, Table 6 below presents the eigenvector for the year 2003 and compares those values with the mean scores of the PISA tests.

Insert Table 6 here

The first point to be noted is that the unadjusted eigenvector has substantially changed between the years 2003 and 2012. Indeed, the difference between the eigenvector components for the years 2003 and 2012 is significantly different from zero at conventional significance levels for 19 out of the 30 OECD

¹⁶This is the case of Italy and Spain, countries that show an internal diversity of educational outcomes (differences between the regions within each country), similar to that of the OECD. This is so in spite of formally having a common educational system.

countries.¹⁷ Most of the countries that performed over the mean in the year 2003 lowered their relative advantage in students' mathematical ability during the period of analysis, with the reduction being largest for the Finland (30%), Netherlands (26%), New Zealand (24%) and Australia (25%). Korea is the exception to the latter rule since it already performed over the mean in 2003 and it improved its relative advantage from 2003 to 2012 by 38%, moving from the 5th position in the 2003 ranking to the first position in 2012. We also find that 11 out of the 15 countries that performed below the mean in 2003 lowered their relative disadvantage during the first decade of the 21th Century. The improvement is particularly relevant for Poland (86%), Portugal (69%), Turkey (65%) and Greece (64%). Poland is the only one of the latter countries that moves to performing over the mean in the year 2012.

As a result of these changes, the coefficient of variation of the unadjusted eigenvector lowered by 11% from 2003 to 2012, attesting that inequality in students' mathematical ability across OECD countries lowered during the first decade of the 21th Century. Additionally, the Kendall's correlation between the unadjusted eigenvectors in 2003 and in 2012 is 0.63, indicating that the ordering of countries in terms of relative performance in mathematical competence has changed from 2003 to 2012. The change is particularly relevant for Poland, that improves its relative position in 18 positions, and for Denmark and Sweden, that lose 12 positions each.

The relationship between the covariate-adjusted eigenvector and the unadjusted one in 2003 is pretty much the same as in 2012, so that we shall not repeat the analysis here. That is, the part of the international differences in relative mathematical performance explained by students' environmental characteristics has remained relatively constant at approximately 50% between the years 2003 and 2012.

Regarding the difference between the covariate-adjusted eigenvector components for the years 2003 and 2012, we find that it is significantly different from zero at conventional significance levels for only 10 out of the 30 OECD countries. In particular, Poland, Portugal and Mexico stand out among the countries that significantly improve their relative position since they increase their component by between 23% and 27%. Conversely, Netherlands lowers its relative advantage once accounting for differences in characteristics by 32% from 2003 to 2012. As a result, the Kendall's correlation between the 2003 and the 2012 covariate-adjusted eigenvectors is of 0.64. The coefficient of variation of

¹⁷We use bootstrap to test whether the difference between a country's eigenvector component in 2003 and in 2012 is statistically significant or not. We implement bootstrap hypothesis testing as follows. For each country, let n_t^c indicate the number of students with level of reading competence c in year t , for $t = \{2003, 2012\}$. For each country and for each c , we merge the samples of the two years into one sample of $(n_c^{2003} + n_c^{2012})$ observations. We draw a bootstrap sample of $(n_c^{2003} + n_c^{2012})$ observations with replacement from the merged sample and we assign the first n_c^{2003} observations to the first year. We then calculate the eigenvectors of the two years and we compute the difference between them. We repeat these steps 1000 times. The p-value is then estimated as the number of times the difference between the eigenvectors coming from bootstrap samples exceeds that observed in the original sample, divided over the number of repetitions.

the covariate-adjusted eigenvector (1) lowered by 16% between the years 2003 and 2012.

Our results remain almost unchanged when we restrict the analysis to the 21 European OECD countries in Table 2. In particular, we find that students' external factors account for slightly more than one half of the differences in relative performance in 2012 and that differences across European OECD countries in students' mathematical ability have decreased from 2003 to 2012 more than they have done in the OECD.¹⁸

3.4 What about differences in reading competence?

We finally tested whether relative performance in reading and mathematical competences evolved in a similar way during the first decade of the 21st century or not. We use the 2009 and 2000 PISA reports that focused on reading competence to analyze country differences in that subject and we restrict the sample to the 27 OECD countries that participated in the four PISA reports. The results are reported in Table 7. First, the results for the mathematical competence are qualitatively identical to those in the preceding tables. Second, we find that the differences in both students' mathematical and reading ability across OECD countries lowered during the first decade of the 21st Century, particularly so in the latter case. The cross-country correlation between the change in relative performance in mathematics and reading is close to 0.7. That is, the countries that improve (lower) their relative position in one of the two competences are likely to be those that improve (lower) their relative position in the other competence analyzed.

Insert Table 7 here

We also find that students' environmental factors account for approximately one half of country differences in relative performance in both mathematics and reading in the years 2003 and 2000, respectively. The relevance of students' familiar and schooling factors in accounting for country differences in relative performance lowered during the period of analysis. Anyway, students' external factors still account for 38% and 46% of country differences in relative performance in reading and mathematics in 2009 and 2012, respectively.

4 Final Remarks

We have presented here a model that combines the one developed in Herrero and Villar (2013), that permits evaluating group performance with categorical data, with inverse probability weighting (IPW) techniques that control for differences in the distribution of the determinants of the outcome variable. We obtain in this way a covariate-adjusted eigenvector that, when compared with the unadjusted one, allows us to estimate the impact of the difference in characteristics over the relative performance.

¹⁸These estimates are available upon request to the authors.

We have applied this methodology to the evaluation of compulsory education in the OECD through PISA 2003 and 2012. We find that differences in mathematical ability across OECD countries would lower by approximately 50% if their endowment of schooling and family characteristics would be that for the OECD average. Equivalently, country differences in reading ability would lower by between 40% and 50%. We have also found that the differences in students' mathematical and reading ability across OECD countries lowered during the first decade of the 21st Century, particularly so in the latter case.

There are two related questions that come to mind when considering this particular application. First, why making an evaluation out of categorical data (the distribution of students across the different levels of competence) when we have all cardinal information that might be required? The reason is that rather than relying on summary statistics (e.g. means and inequality measures) as it is the case in most of the cardinal approaches, we are able to deal with discretized versions of the whole distributions in a relatively simple way and so to extract more information. Second, why to use just five levels of competence rather than richer distribution profiles (e.g. using percentiles)? The answer here is twofold. On the one hand, a small number of levels permits a richer set of covariates. On the other hand, in this particular case, those levels are given externally so that there is less arbitrariness in deciding the clusters by the analysts.

Acknowledgements: We thank the editor and an anonymous referee for their comments and suggestions that led to considerable improvements of the paper. Thanks are also due to Dirk Van de gaer, X. Ramos, F. Goerlich and the participants in different seminars for helpful comments. This paper was written while the first author was visiting the IAE, under the Salvador de Madariaga project PR2010-0124. This paper is framed within the projects SEJ2007-62656, SEJ-6882/ECON and ECO2010-21706.

5 References

- Arneson, R.J., 1989. Equality and equality of opportunity for welfare. *Philosophical Studies* 56, 77-93.
- Busso, M., DiNardo, J., McCrary, J., 2012. New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators. IZA Discussion Paper 3998.
- Calo-Blanco, A., Villar, A., 2010. Quality of Education and Equality of Opportunity in Spain. Lessons from PISA, Fundación BBVA, working paper 6-2010.
- Calo-Blanco, A., García-Pérez, J.I., 2013. On the Welfare Loss Caused by Inequality of Opportunity. *Journal of Economic Inequality*. DOI: 10.1007/s10888-013-9247-6.
- Checchi, D., Peragine, V., 2010. Inequality of opportunity in Italy. *Journal of Economic Inequality* 8, 429-450.
- Ciccone, A., Garcia-Fontes, W., 2012. The quality of the Catalan and Spanish education systems: A perspective from PISA. IESE Research Papers D/810, IESE Business School.
- Cohen, G.A., 1989. On the currency of egalitarian justice. *Ethics*, 99(4), 906-944.
- Crump, R.K., Hotz, V.J., Imbens, G.W., Mitnik, O.A., 2012. Dealing with limited overlap in estimation of average treatment effects. *Biometrika* 96(1), 187-199.
- Echenique, F., Fryer, R.G., 2005. On the measurement of segregation. Working Paper 11258, NBER.
- Fleurbaey, M., 2008. Fairness, responsibility and welfare. Oxford University Press, Oxford.
- Hanushek, E.A., Woessmann, L., 2011. The Economics of International Differences in Educational Achievement. In E.A. Hanushek, S. Machin, L. Woessmann (eds.), *Handbook of the Economics of Education*, Vol. 3, pp. 89-200, Amsterdam: North Holland.
- Herrero, C., Villar, A., 2013. On the Comparison of Group Performance with Categorical Data. *PLoS ONE* 8(12): e84784. doi:10.1371/journal.pone.0084784
- Lefranc, A., Pistolesi, N., Trannoy, A., 2008. Inequality of opportunities vs. inequality of outcomes: Are western societies all alike?. *Review of Income and Wealth* 54, 513-546.
- Lefranc, A., Pistolesi, N., Trannoy, A., 2012. Equality of opportunity and luck: definitions and testable conditions, with an application to income in France. *Journal of Public Economics* 93, 1189-1207.
- Liebertson, S. 1976. Rank-Sum comparisons between Groups. *Sociological Methodology* 7, 276-291.
- OECD, 2012. PISA Data Analysis Manual: SPSS. Second Edition. OECD, Paris.
- OECD, 2010a. PISA 2012 Results: What Students Know and Can Do. Student Performance in mathematical, Mathematics and Science (Volume I).

OECD, Paris.

OECD, 2010b. PISA 2012 Results: Overcoming Social Background. Equity in learning opportunities and outcomes (Volume II). OECD, Paris.

Peragine, V., Serlenga, L., 2008. Higher education and equality of opportunity in Italy. *Research on Economic Inequality* 16, 67-97.

Roemer, J.E., 1993. A pragmatic theory of responsibility for the egalitarian planner. *Philosophy and Public Affairs* 22(2), 146-166.

Roemer, J.E., 1998. *Equality of opportunity*. Harvard University Press, New York.

Rosenbaum, P., Rubin, D.B., 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70, 41-55.

Villar, A., 2012. The Educational Development Index. Working Paper Econ n° 1206, Pablo de Olavide University.

Table 1. Levels of mathematical competence. PISA 2012.

Level of competence		Score range
Level 6		>698
Level 5		626-698
Level 4		553-626
Level 3		480-553
Level 2		407-480
Level 1	Level 1a	335-407
	Level 1b	262-335

Table 2. Share of students at each proficiency level on the mathematical scale. OECD countries. PISA 2012.

Country	Level 1 (or below)	Level 2	Level 3	Level 4	Level 5	Sample size
Australia	0.149	0.251	0.291	0.206	0.100	9,996
Austria	0.120	0.240	0.304	0.239	0.095	3,444
Belgium	0.093	0.175	0.272	0.277	0.180	5,923
Canada	0.095	0.246	0.327	0.233	0.096	16,331
Czech Republic	0.112	0.208	0.269	0.264	0.145	4,108
Denmark	0.142	0.278	0.330	0.197	0.051	5,155
Finland	0.099	0.227	0.336	0.236	0.101	6,460
France	0.128	0.238	0.300	0.233	0.099	3,124
Germany	0.095	0.171	0.281	0.294	0.156	2,614
Greece	0.259	0.329	0.267	0.119	0.023	3,914
Hungary	0.167	0.298	0.298	0.171	0.064	3,527
Iceland	0.146	0.263	0.318	0.204	0.068	2,511
Ireland	0.108	0.262	0.337	0.220	0.072	3,795
Italy	0.153	0.267	0.314	0.197	0.067	24,041
Japan	0.070	0.176	0.306	0.267	0.179	4,613
Korea	0.063	0.152	0.247	0.277	0.260	3,797
Luxembourg	0.188	0.262	0.287	0.195	0.065	3,278
Mexico	0.427	0.368	0.166	0.034	0.003	24,159
Netherlands	0.094	0.199	0.286	0.280	0.140	3,339
New Zealand	0.136	0.235	0.281	0.219	0.127	2,752
Norway	0.148	0.272	0.320	0.196	0.063	3,413
Poland	0.080	0.234	0.302	0.249	0.133	3,449
Portugal	0.172	0.260	0.300	0.194	0.071	4,245
Slovak Republic	0.187	0.250	0.268	0.204	0.088	3,470
Spain	0.135	0.258	0.327	0.219	0.059	19,461
Sweden	0.176	0.282	0.309	0.181	0.049	3,446
Switzerland	0.094	0.209	0.304	0.258	0.133	8,363
Turkey	0.321	0.314	0.196	0.120	0.047	3,457
United Kingdom	0.137	0.272	0.314	0.199	0.076	8,981
United States	0.180	0.297	0.291	0.170	0.059	3,633

Table 3. Eigenvectors and mean score. OECD, 2012. Mathematical competence.

Country	Unadjusted	Mean score ^a	Covariate-adjusted	
			(1)	(2)
Australia	0.931***	1.017	0.915	0.932
Austria	1.042	1.031	1.091	0.945
Belgium	1.444**	1.081	1.135**	1.229
Canada	1.079	1.038	0.985	1.001
Czech Republic	1.279	1.057	1.193	1.179
Denmark	0.808***	0.998	0.946	0.972
Finland	1.128***	1.040	0.930	0.978
France	1.052	1.027	1.047	1.039
Germany	1.428***	1.076	0.928	0.929
Greece	0.452***	0.930	0.735	0.742
Hungary	0.762	0.987	0.919	0.950
Iceland	0.901	1.004	0.773	0.738
Ireland	0.937	1.020	0.954	1.016
Italy	0.837	1.000	0.999	0.995
Japan	1.667**	1.087	1.089	1.127
Korea	2.080***	1.125	1.048	1.010
Luxembourg	0.845**	0.989	0.988***	0.996
Mexico	0.209***	0.851	0.653***	0.633
Netherlands	1.363***	1.066	1.153***	1.174
New Zealand	1.088***	1.035	0.909**	0.917
Norway	0.813	1.000	0.766	0.767
Poland	1.242***	1.060	1.606***	1.647
Portugal	0.760***	0.995	1.180**	1.125
Slovak Republic	0.813	0.997	1.284	1.191
Spain	0.884**	1.007	0.891**	0.905
Sweden	0.718***	0.983	0.815	0.841
Switzerland	1.358***	1.059	1.299	1.269
Turkey	0.455***	0.924	0.867**	0.871
United Kingdom	0.888**	1.007	0.896	0.927
United States	0.737	0.982	1.005***	0.954
Coeff. of variation	0.370	0.055	0.190	0.190

Notes: ^a The mean is set equal to one. The symbols ** and *** indicate that the difference between the component and that for the same country and eigenvector (adjusted or unadjusted) in the year 2003 is significant at the 5% and 1% significance level, respectively. We use bootstrap hypothesis testing. The covariate-adjusted eigenvector in (1) controls for students' sex, their parents' employment status and educational level, whether at least one of the parents was born in a country different from their current country of residence, whether there are at least one hundred books at home, the percentage of the school total funding for a typical school year that comes from the government and whether the assessments of students are used to make judgements about teachers' effectiveness. The eigenvector in (2) further controls for whether the school is located in a city with over one million people, whether the principal, the department head or the teachers have the main responsibility for hiring teachers or not, the share of teachers fully certified by the appropriate authority in the school, whether students are grouped by ability, whether achievement data are used to evaluate the principal's or the teachers' performance and whether the school monitors the practice of teachers.

Table 4. Students' personal and familiar characteristics. Descriptive statistics. PISA 2012.

Country	Father		Mother		Foreign		Mother's educ.		Father's education		100 books ^b
	Women	works	works	parent ^a	Secondary	Tertiary	Secondary	Tertiary	Secondary	Tertiary	
Australia	48.9	90.0	76.1	35.0	53.3	43.4	56.9	37.0	48.4		
Austria	50.5	93.6	80.1	22.8	70.2	27.8	56.4	42.1	42.3		
Belgium	50.8	91.0	79.7	26.0	40.2	56.1	46.1	50.0	40.7		
Canada	51.2	91.7	80.9	27.3	35.4	63.4	44.1	54.0	45.1		
Czech Republic	50.3	93.9	84.8	10.6	71.2	28.5	71.3	28.3	47.1		
Denmark	51.4	84.9	78.6	31.3	38.8	57.1	52.0	44.6	36.0		
Finland	50.7	87.2	84.3	23.0	25.4	72.0	32.4	63.1	43.3		
France	52.2	90.3	80.0	24.7	50.4	47.6	55.6	42.4	39.9		
Germany	50.5	94.3	80.4	15.4	68.0	29.9	52.1	45.9	54.8		
Greece	51.7	81.6	58.1	18.1	51.0	43.0	49.8	42.1	37.1		
Hungary	52.7	86.9	76.0	5.5	61.9	37.7	68.9	30.5	51.7		
Iceland	50.4	94.1	85.1	11.2	43.6	56.4	51.5	48.5	56.8		
Ireland	50.0	81.6	63.4	27.3	52.8	44.2	54.2	40.1	42.6		
Italy	49.8	92.0	65.9	14.0	68.7	28.6	70.5	26.0	40.5		
Japan	47.6	97.0	77.1	1.5	50.5	49.5	49.7	50.3	43.1		
Korea	45.2	90.5	58.7	0.4	56.1	42.5	44.8	53.6	62.6		
Luxembourg	48.7	90.2	72.3	65.2	45.1	38.4	44.1	40.3	50.5		
Mexico	52.2	85.6	40.4	3.1	42.1	28.3	40.4	31.6	8.7		
Netherlands	49.3	91.8	78.6	18.2	49.9	45.5	44.5	49.5	35.4		
New Zealand	51.7	92.3	77.5	43.0	54.1	44.0	58.8	37.3	48.1		
Norway	49.4	92.7	86.4	18.0	39.7	59.0	49.6	49.5	49.9		
Poland	52.9	87.6	71.4	0.9	75.3	24.6	81.3	18.4	38.1		
Portugal	50.6	86.2	75.3	19.7	47.7	23.2	44.1	19.7	30.3		
Slovak Republic	46.7	87.2	77.4	5.2	74.1	25.3	76.5	23.0	31.2		
Spain	50.6	86.2	69.1	15.6	49.0	41.6	46.3	41.1	50.2		
Sweden	51.5	93.4	89.5	25.2	35.7	62.4	45.8	51.4	47.8		
Switzerland	50.6	94.1	75.4	42.9	57.8	37.8	47.5	48.5	37.7		
Turkey	51.6	71.2	14.8	2.8	27.3	11.5	37.9	23.9	19.9		
United Kingdom	50.3	88.9	77.3	18.4	50.3	49.0	55.1	43.8	40.1		
United States	49.6	85.9	73.7	29.6	42.7	52.1	49.7	44.3	30.8		

Notes: ^a Informs on whether at least one of the student's parents was born in a country different from their current country of residence. ^b Indicates whether there are at least one hundred books at home or not.

Table 5. Schooling characteristics. Descriptive statistics. PISA 2012.

Country	Large city ^c	Fund. govern. ^d	Evaluate teachers ^e	Control ^f	Track ^g	Monitor ^h	Certified teachers ⁱ
Australia	39.0	68.1	96.9	64.8	80.2	90.8	86.3
Austria	34.2	69.9	96.8	67.9	69.7	91.6	77.6
Belgium	33.4	72.3	95.3	71.2	73.6	92.3	77.9
Canada	37.3	67.8	97.4	62.0	79.8	91.0	88.0
Czech Republic	31.7	69.0	95.5	71.5	72.1	90.7	81.6
Denmark	31.1	71.2	96.4	72.5	72.8	90.0	78.5
Finland	32.0	70.5	95.7	70.6	72.4	90.3	79.4
France	29.9	71.9	96.8	71.2	69.8	93.3	76.6
Germany	29.5	73.0	96.2	69.9	70.8	92.8	76.9
Greece	32.7	70.8	96.0	69.3	69.5	91.4	78.9
Hungary	31.5	70.1	98.2	67.7	72.1	90.1	79.5
Iceland	26.6	78.9	97.9	78.0	76.9	93.5	79.3
Ireland	33.9	70.7	96.6	70.7	70.4	91.7	78.0
Italy	36.5	59.3	97.7	51.9	76.5	91.5	87.4
Japan	32.8	69.8	96.8	70.3	71.2	91.5	77.1
Korea	29.4	71.1	97.3	72.2	72.9	94.1	79.0
Luxembourg	32.0	70.1	100.0	65.4	76.3	96.1	89.2
Mexico	39.6	53.4	98.2	52.8	75.4	92.8	87.1
Netherlands	34.0	69.6	97.2	69.9	71.3	91.9	78.4
New Zealand	32.2	70.5	97.1	71.4	73.3	92.1	80.0
Norway	30.7	71.0	97.3	69.1	71.7	91.5	77.7
Poland	34.7	72.7	96.9	69.1	69.4	91.6	81.0
Portugal	33.7	69.5	97.0	69.5	71.2	91.9	77.9
Slovak Republic	34.4	73.4	96.2	75.2	67.5	92.1	73.4
Spain	37.8	66.0	97.6	61.1	80.8	89.9	88.0
Sweden	32.8	71.1	96.4	69.8	72.4	92.2	78.6
Switzerland	33.4	73.0	96.3	70.6	74.6	88.8	82.1
Turkey	32.5	72.0	98.0	70.3	71.5	93.3	77.5
United Kingdom	36.2	72.0	97.2	70.9	77.8	90.6	82.5
United States	29.2	70.7	97.7	73.0	72.3	92.4	78.5

Notes: ^c Indicates whether the school is located in a city with over one million people or not. ^d Informs on the percentage of the school total funding for a typical school year that comes from the government. ^e Indicates whether the assessments of students are used to make judgements about teachers' effectiveness or not. ^f Indicates whether the principal, the department head or the teachers have the main responsibility for hiring teachers or not. ^g Indicates whether students are grouped by ability or not. ^h Indicates whether the school monitors the practice of teachers. ⁱ Informs on the share of teachers fully certified by the appropriate authority in the school.

Table 6. Eigenvectors and mean scores. OECD countries, 2003. Mathematical competence.

Countries	Unadjusted eigenvector	Mean score ^a	Covariate-adj. eigenvector
Australia	1.238	1.046	0.892
Austria	1.010	1.025	1.182
Belgium	1.704	1.102	1.272
Canada	1.150	1.046	0.964
Czech Republic	1.437	1.077	1.102
Denmark	1.109	1.034	0.859
Finland	1.616	1.082	1.001
France	1.073	1.040	1.162
Germany	1.126	1.027	0.938
Greece	0.275	0.880	0.716
Hungary	0.692	0.967	1.001
Iceland	0.990	1.029	0.850
Ireland	0.892	1.009	1.060
Italy	0.799	0.978	0.926
Japan	1.514	1.056	0.999
Korea	1.502	1.080	1.057
Luxembourg	0.735	1.018	1.228
Mexico	0.156	0.817	0.521
Netherlands	1.835	1.100	1.684
New Zealand	1.426	1.063	1.028
Norway	0.833	0.997	0.695
Poland	0.667	0.967	1.270
Portugal	0.451	0.932	0.962
Slovak Republic	0.882	1.013	1.138
Spain	0.761	0.995	0.796
Sweden	1.023	1.025	0.811
Switzerland	1.064	1.039	1.365
Turkey	0.277	0.860	0.774
United Kingdom	1.089	1.041	0.912
United States	0.675	0.970	0.834
Coeff. of variation	0.416	0.067	0.226

Notes: ^a The mean is set equal to one. The covariate-adjusted eigenvector controls for students' sex, their parents' employment status and educational level, whether at least one of the parents was born in a country different from their current country of residence, whether there are at least one hundred books at home, the percentage of the school total funding for a typical school year that comes from the government and whether the assessments of students are used to make judgements about teachers' effectiveness.

Table 7. Eigenvectors. OECD countries. Mathematical (2012 and 2003) and reading (2009 and 2000) competences.

Country	Mathematics				Reading			
	2012		2003		2009		2000	
	Unadj.	Adj.	Unadj.	Adj.	Unadj.	Adj.	Unadj.	Adj.
Australia	0.940***	0.909	1.213	0.876	1.173	0.956	1.345	1.000
Austria	1.049	1.071	1.000	1.241	0.684**	0.825***	0.882	1.307
Belgium	1.496	1.122**	1.657	1.360	1.343	0.954	1.245	1.042
Canada	1.069	1.004	1.154	0.994	1.138	0.983	1.283	0.987
Czech Republic	1.218**	1.158	1.482	1.109	0.929	1.212***	0.851	0.801
Denmark	0.824***	0.977	1.062	0.891	0.743	0.722	0.885	0.849
Finland	1.101***	0.979	1.631	1.006	1.502***	1.234	1.902	1.303
France	1.048	1.053	1.125	1.204	1.076	0.930	0.947	1.100
Germany	1.475***	0.949	1.063	0.896	1.171***	0.979	0.897	1.144
Greece	0.443***	0.728	0.302	0.761	0.760***	0.792	0.591	0.845
Hungary	0.703	0.974	0.660	0.885	0.852	0.995***	0.655	0.703
Iceland	0.860***	0.722**	1.066	0.885	0.969	0.856	0.985	0.868
Ireland	0.965***	0.928	0.835	1.043	0.979***	1.012	1.382	1.028
Italy	0.820	1.011	0.764	0.913	0.801	1.031***	0.722	0.799
Korea	2.097***	1.005	1.513	1.135	1.844***	1.394***	1.198	1.084
Luxembourg	0.982***	1.050	0.740	1.243	0.747***	1.029***	0.410	0.836
Mexico	0.207***	0.649***	0.147	0.479	0.275	0.583***	0.270	0.761
Netherlands	1.343**	1.154	1.625	1.338	1.302***	1.230	1.825	1.319
New Zealand	1.068**	0.908	1.367	0.976	1.609	1.669***	1.394	1.187
Norway	0.843	0.781	0.783	0.718	0.958	0.736	1.021	0.834
Poland	1.223***	1.624***	0.635	1.239	1.025***	1.178***	0.572	0.803
Portugal	0.811***	1.238***	0.456	1.037	0.719	0.933***	0.630	1.238
Spain	0.860**	0.897	0.740	0.802	0.738	0.921	0.778	0.993
Sweden	0.694***	0.825	1.014	0.825	0.964**	0.907	1.167	0.873
Switzerland	1.312	1.292	1.173	1.336	0.878	1.000	0.850	1.150
United Kingdom	0.870***	0.901	1.144	0.961	0.904***	0.908***	1.323	1.216
United States	0.681	1.093***	0.648	0.847	0.916	1.031	0.992	0.931
Coeff. of var.	0.356	0.194	0.392	0.209	0.316	0.197	0.381	0.182

Notes: The symbols ** and *** indicate that the difference between the component and that for the same country, subject and eigenvector (adjusted or unadjusted) in the preceding year is significant at the 5% and 1% significance level, respectively. We use bootstrap hypothesis testing. The covariate-adjusted eigenvector controls for students' sex, their parents' employment status and educational level, whether at least one of the parents was born in a country different from their current country of residence, whether there are at least one hundred books at home, the percentage of the school total funding for a typical school year that comes from the government and whether the assessments of students are used to make judgements about teachers' effectiveness.

FOR THE REFEREES

Table R.1. Share of students at each proficiency level on the mathematical scale. OECD countries. PISA 2003.

Country	Level 1 (or below)	Level 2	Level 3	Level 4	Level 5	Sample size
Australia	0.104	0.200	0.284	0.265	0.144	7,195
Austria	0.105	0.221	0.319	0.251	0.101	2,849
Belgium	0.069	0.149	0.241	0.285	0.252	4,752
Canada	0.072	0.215	0.327	0.269	0.115	11,120
Czech Republic	0.076	0.194	0.263	0.262	0.202	3,703
Denmark	0.089	0.222	0.320	0.257	0.109	2,403
Finland	0.040	0.163	0.335	0.298	0.162	3,391
France	0.092	0.209	0.311	0.271	0.115	2,335
Germany	0.124	0.213	0.300	0.241	0.120	2,509
Greece	0.354	0.301	0.220	0.104	0.019	2,563
Hungary	0.183	0.294	0.283	0.183	0.054	3,022
Iceland	0.106	0.222	0.304	0.261	0.105	1,758
Ireland	0.109	0.248	0.346	0.216	0.079	2,276
Italy	0.167	0.266	0.317	0.185	0.063	6,883
Japan	0.088	0.192	0.291	0.278	0.149	3,007
Korea	0.065	0.176	0.276	0.300	0.180	3,334
Luxembourg	0.124	0.212	0.321	0.252	0.088	1,558
Mexico	0.482	0.338	0.148	0.028	0.002	11,273
Netherlands	0.042	0.162	0.279	0.294	0.221	2,192
New Zealand	0.092	0.186	0.288	0.259	0.173	2,182
Norway	0.148	0.254	0.301	0.218	0.076	2,329
Poland	0.174	0.300	0.291	0.174	0.058	2,774
Portugal	0.232	0.320	0.280	0.135	0.031	2,740
Slovak Republic	0.126	0.248	0.296	0.225	0.103	4,493
Spain	0.127	0.258	0.338	0.221	0.054	5,997
Sweden	0.103	0.238	0.308	0.243	0.106	2,662
Switzerland	0.110	0.208	0.291	0.254	0.133	4,602
Turkey	0.431	0.282	0.163	0.076	0.044	2,860
United Kingdom	0.091	0.214	0.308	0.263	0.122	4,897
United States	0.188	0.274	0.287	0.182	0.067	3,216

Table R2. Students' personal and familiar characteristics. Descriptive statistics. PISA 2003.

Country	Women	Father		Foreign parent ^a	Mother's educ.		Father's education		100 books ^b
		works	Mother works		Secondary	Tertiary	Secondary	Tertiary	
Australia	49.2	13.2	71.7	40.8	55.1	40.6	53.4	42.3	59.5
Austria	53.0	21.9	69.3	16.5	75.4	22.3	59.7	38.8	45.9
Belgium	48.9	25.0	68.4	22.2	43.9	48.9	45.2	48.5	44.0
Canada	51.1	16.1	78.3	19.3	50.7	47.4	54.5	41.6	53.6
Czech Republic	48.5	5.9	85.5	7.4	78.9	20.5	75.0	24.5	66.1
Denmark	49.8	19.5	83.6	10.9	46.4	50.8	58.1	38.2	47.1
Finland	51.9	16.6	83.3	4.8	38.0	57.0	41.4	49.8	47.4
France	51.8	21.0	73.1	25.4	65.0	28.6	61.9	32.0	44.3
Germany	51.4	17.7	72.8	18.4	69.7	22.4	57.5	35.1	52.3
Greece	50.4	29.8	51.4	14.6	57.6	26.0	50.7	31.9	33.9
Hungary	45.5	14.7	72.5	3.4	75.5	23.9	78.2	21.3	58.0
Iceland	49.8	8.0	85.8	7.1	66.2	27.6	65.2	30.3	64.9
Ireland	53.0	24.0	63.0	15.2	61.5	28.8	57.4	28.4	43.5
Italy	51.7	22.5	64.6	6.6	68.8	23.8	68.8	23.5	43.3
Japan	52.7	22.0	74.1	0.6	52.2	45.4	47.7	45.5	45.3
Korea	39.3	16.2	45.7	0.1	64.9	22.2	55.7	34.6	49.7
Luxembourg	46.7	21.2	61.2	51.2	35.9	45.2	36.6	47.1	52.6
Mexico	53.9	65.8	36.5	2.7	37.9	23.7	39.0	29.3	10.6
Netherlands	49.1	27.7	72.5	18.2	58.7	30.5	49.9	40.1	44.8
New Zealand	46.8	16.5	78.3	38.6	56.1	37.3	60.7	30.2	54.9
Norway	49.6	11.9	83.2	12.0	46.6	51.5	48.7	49.0	64.2
Poland	50.6	25.6	58.7	0.3	82.3	16.2	82.5	13.8	40.8
Portugal	53.1	33.4	70.5	11.3	30.3	21.9	34.3	19.8	31.8
Slovak Republic	49.1	14.5	81.2	6.6	79.2	19.7	75.3	23.8	46.2
Spain	51.6	13.5	61.4	5.5	41.4	29.2	36.4	34.8	53.8
Sweden	49.6	11.2	85.1	16.9	44.4	52.9	49.1	45.1	61.2
Switzerland	48.7	24.5	66.8	38.2	70.5	22.8	56.8	37.9	43.4
Turkey	43.1	48.6	11.1	2.2	30.1	10.1	43.5	21.1	19.4
United Kingdom	53.2	22.4	75.0	17.3	60.7	36.3	62.4	31.8	47.2
United States	49.6	26.9	76.2	21.3	59.0	38.1	61.1	35.3	41.8

Notes: ^a Indicates whether at least one of the parents was born in a country different from their current country of residence or not. ^b Indicates whether there are at least one hundred books at home or not.

Table R.3. Schooling characteristics. Descriptive statistics. PISA 2003.

Country	Large city ^c	Fund. govern. ^d	Evaluate teachers ^e	Control ^f	Track ^g	Monitor ^h	Certified teachers ⁱ
Australia	32.4	67.3	97.1	69.2	74.1	88.3	62.3
Austria	29.3	71.5	95.9	69.7	65.7	87.6	71.5
Belgium	32.5	68.4	96.8	70.3	70.4	89.3	61.9
Canada	30.7	47.5	97.6	75.7	85.9	88.0	79.1
Czech Republic	35.8	66.7	96.6	72.6	73.4	89.4	69.6
Denmark	33.8	69.3	94.9	73.6	71.3	86.8	70.5
Finland	32.5	68.6	94.8	70.8	69.7	84.3	69.7
France	36.5	71.4	97.4	67.3	69.6	83.5	72.6
Germany	33.4	69.8	96.0	72.0	69.4	86.6	69.3
Greece	36.3	71.1	94.8	69.5	77.4	84.3	67.8
Hungary	33.6	67.2	96.2	66.2	69.8	87.8	63.1
Iceland	34.9	75.4	97.3	72.0	68.9	87.9	59.0
Ireland	37.9	68.5	98.6	64.8	71.0	83.8	68.3
Italy	34.1	66.1	96.3	67.7	76.5	88.4	61.6
Japan	36.0	69.8	97.6	64.8	70.8	84.2	69.2
Korea	38.3	71.1	96.5	66.9	69.5	84.4	70.3
Luxembourg	29.0	76.1	99.7	79.5	74.2	91.6	57.9
Mexico	31.1	48.4	97.1	76.6	84.9	87.2	77.7
Netherlands	37.2	71.1	96.6	66.7	71.1	85.4	71.0
New Zealand	36.2	71.8	97.2	68.3	72.6	83.6	71.7
Norway	32.6	70.5	95.5	68.8	70.9	84.5	71.1
Poland	37.0	69.9	97.4	67.0	71.5	84.3	72.0
Portugal	37.9	70.4	96.3	66.7	70.8	84.6	70.7
Slovak Republic	31.1	68.1	96.0	72.2	75.6	88.1	62.8
Spain	32.6	67.0	96.9	69.4	76.4	88.4	61.2
Sweden	33.3	70.1	95.7	68.5	69.5	83.8	72.7
Switzerland	30.8	64.6	95.8	70.7	78.1	88.9	62.5
Turkey	34.3	70.7	97.2	69.4	71.3	84.4	68.9
United Kingdom	32.4	67.3	97.0	68.3	76.3	88.9	62.8
United States	35.5	67.3	96.1	69.2	74.0	89.4	63.4

Notes: ^c Indicates whether the school is located in a city with over one million people or not. ^d Informs on the percentage of the school total funding for a typical school year that comes from the government. ^e Indicates whether the assessments of students are used to make judgements about teachers' effectiveness or not. ^f Indicates whether the principal, the department head or the teachers have the main responsibility for hiring teachers or not. ^g Indicates whether students are grouped by ability or not. ^h Indicates whether the school monitors the practice of teachers. ⁱ Informs on the share of teachers fully certified by the appropriate authority in the school.

Table R.4. Share of students at each proficiency level on the reading scale.
OECD countries. PISA 2009.

Country	Level 1 (or below)	Level 2	Level 3	Level 4	Level 5	Sample size
Australia	0.117	0.204	0.296	0.260	0.123	9,975
Austria	0.215	0.249	0.288	0.203	0.045	4,571
Belgium	0.098	0.183	0.280	0.316	0.123	5,731
Canada	0.107	0.216	0.305	0.264	0.107	16,782
Czech Republic	0.145	0.235	0.277	0.252	0.091	4,283
Denmark	0.155	0.272	0.325	0.213	0.034	3,991
Finland	0.073	0.165	0.302	0.329	0.131	4,372
France	0.134	0.207	0.302	0.255	0.102	2,787
Germany	0.118	0.210	0.305	0.280	0.088	2,853
Greece	0.165	0.266	0.316	0.205	0.048	3,676
Hungary	0.142	0.254	0.346	0.217	0.040	3,203
Iceland	0.143	0.215	0.332	0.231	0.079	2,581
Ireland	0.132	0.228	0.329	0.244	0.067	2,596
Italy	0.163	0.241	0.315	0.229	0.051	22,617
Korea	0.036	0.135	0.323	0.376	0.130	3,554
Luxembourg	0.198	0.239	0.293	0.210	0.059	3,053
Mexico	0.337	0.373	0.236	0.052	0.002	21,134
Netherlands	0.082	0.221	0.311	0.284	0.102	3,304
New Zealand	0.093	0.168	0.271	0.286	0.182	2,885
Norway	0.126	0.227	0.326	0.243	0.079	3,299
Poland	0.104	0.234	0.328	0.260	0.075	3,370
Portugal	0.163	0.257	0.332	0.210	0.037	4,748
Spain	0.159	0.250	0.352	0.206	0.033	18,947
Sweden	0.134	0.221	0.327	0.225	0.094	3,132
Switzerland	0.138	0.238	0.325	0.234	0.065	8,502
United Kingdom	0.137	0.241	0.310	0.229	0.084	8,288
United States	0.148	0.247	0.287	0.229	0.089	3,569
OECD total	0.161	0.247	0.306	0.219	0.066	185,245
OECD average	0.145	0.234	0.307	0.236	0.078	6,388

Table R.5. Students' personal and familiar characteristics. Descriptive statistics. PISA 2009.

Country	Mother works		Father works		Foreign parent ^a		Mother's educ.		Fathers' educ.		100 books ^b
	Women	works	works	works	parent ^a	parent ^a	Secondary	Tertiary	Secondary	Tertiary	
Australia	51.4	76.8	90.5	39.1	56.2	42.9	59.1	39.7	52.3		
Austria	51.7	76.2	93.2	20.5	73.8	24.6	55.2	43.8	42.9		
Belgium	49.7	77.1	90.9	25.4	41.7	54.4	46.4	49.1	42.6		
Canada	51.3	81.4	90.9	24.8	38.6	60.3	46.8	51.1	48.6		
Czech Republic	48.0	84.0	94.3	8.3	73.9	25.9	72.5	27.2	49.3		
Denmark	51.6	80.1	87.1	29.2	43.1	52.1	56.0	40.7	37.7		
Finland	51.7	87.4	89.0	7.1	24.1	72.6	28.3	66.6	52.3		
France	53.0	77.3	90.7	23.6	55.4	42.1	58.8	38.9	41.3		
Germany	50.4	75.9	91.1	19.7	69.7	27.4	55.9	42.0	50.0		
Greece	51.8	63.1	89.4	15.6	54.4	38.7	51.1	40.3	39.1		
Hungary	50.0	75.6	86.0	5.1	69.2	30.3	74.1	25.4	52.0		
Iceland	52.3	85.5	92.4	9.1	55.1	44.0	57.0	42.0	58.5		
Ireland	49.8	67.8	86.6	23.7	57.1	39.5	56.8	36.2	43.1		
Italy	50.1	65.3	93.1	12.0	72.4	24.4	71.9	23.6	41.2		
Korea	45.9	52.3	88.1	0.2	63.2	34.3	48.6	48.6	57.2		
Luxembourg	52.8	67.2	92.1	55.5	47.8	38.0	44.9	42.2	55.7		
Mexico	52.8	39.5	86.7	3.0	40.6	28.3	38.5	33.0	11.2		
Netherlands	51.4	78.3	91.8	18.9	53.5	40.7	47.3	46.1	38.6		
New Zealand	51.4	79.0	91.2	40.8	56.4	41.1	60.6	36.7	50.6		
Norway	49.1	87.6	92.8	14.4	38.9	60.3	44.7	54.1	54.4		
Poland	50.3	69.3	85.7	0.6	79.2	20.6	84.9	14.8	37.9		
Portugal	52.2	76.3	90.1	19.4	45.5	21.9	43.9	18.9	31.2		
Spain	49.8	68.4	90.1	14.5	50.5	37.0	46.3	38.7	49.5		
Sweden	50.4	87.4	92.0	21.3	35.4	62.9	46.3	50.6	54.7		
Switzerland	49.4	72.4	93.9	40.4	59.5	36.3	48.1	48.6	40.8		
United Kingdom	51.1	77.3	88.9	13.1	51.3	47.6	55.4	43.1	42.3		
United States	49.4	73.7	85.1	26.4	48.0	47.8	54.0	41.1	35.0		

Notes: ^a Informs on the percentage of the school total funding for a typical school year that comes from the government. ^b Indicates whether there are at least one hundred books at home or not.

Table R.6. Schooling characteristics. Descriptive statistics. PISA 2009.

Country	Fund govern. ^c	Evaluate teachers ^d
Australia	57.7	38.6
Austria	60.6	40.5
Belgium	60.5	43.8
Canada	45.1	38.6
Czech Republic	60.6	40.7
Denmark	61.8	39.9
Finland	62.3	45.9
France	56.0	53.5
Germany	60.4	42.7
Greece	59.0	46.6
Hungary	59.1	47.6
Iceland	50.6	43.2
Ireland	58.9	51.7
Italy	45.7	40.5
Korea	41.9	51.0
Luxembourg	43.1	50.2
Mexico	43.1	45.3
Netherlands	42.2	48.6
New Zealand	42.6	51.8
Norway	42.1	46.8
Poland	42.2	49.1
Portugal	42.1	42.2
Spain	43.5	41.1
Sweden	41.8	49.1
Switzerland	43.5	41.2
United Kingdom	43.8	40.3
United States	42.6	52.3

Notes: ^c Informs on the percentage of the school total funding for a typical school year that comes from the government. ^d Indicates whether the assessments of students are used to make judgements about teachers' effectiveness or not.

Table R.7. Share of students at each proficiency level on the reading scale.
OECD countries. PISA 2000.

Country	Level 1 (or below)	Level 2	Level 3	Level 4	Level 5	Sample size
Australia	0.121	0.184	0.256	0.260	0.179	4,866
Austria	0.155	0.233	0.299	0.238	0.074	4,390
Belgium	0.143	0.157	0.270	0.303	0.126	6,206
Canada	0.105	0.195	0.288	0.269	0.143	28,006
Czech Republic	0.148	0.250	0.311	0.219	0.072	5,110
Denmark	0.156	0.223	0.316	0.230	0.075	3,956
Finland	0.056	0.138	0.287	0.338	0.181	4,617
France	0.139	0.226	0.311	0.247	0.078	4,396
Germany	0.176	0.217	0.283	0.227	0.097	4,751
Greece	0.234	0.266	0.286	0.174	0.040	4,353
Hungary	0.200	0.262	0.308	0.191	0.038	4,521
Iceland	0.133	0.217	0.315	0.252	0.082	3,211
Ireland	0.096	0.171	0.308	0.285	0.141	3,631
Italy	0.173	0.260	0.319	0.202	0.045	4,785
Korea	0.055	0.193	0.405	0.309	0.037	4,583
Luxembourg	0.305	0.286	0.276	0.122	0.012	3,194
Mexico	0.389	0.324	0.215	0.066	0.006	4,005
Netherlands	0.063	0.138	0.287	0.346	0.166	2,359
New Zealand	0.120	0.174	0.252	0.272	0.181	3,396
Norway	0.158	0.197	0.286	0.248	0.111	3,922
Poland	0.248	0.260	0.284	0.166	0.042	3,101
Portugal	0.222	0.251	0.299	0.190	0.037	4,357
Spain	0.146	0.254	0.341	0.225	0.033	5,791
Sweden	0.109	0.198	0.317	0.267	0.109	4,156
Switzerland	0.178	0.223	0.294	0.220	0.085	5,717
United Kingdom	0.111	0.187	0.277	0.263	0.162	8,524
United States	0.156	0.210	0.288	0.233	0.114	2,914
OECD total	0.158	0.218	0.291	0.235	0.097	150,835
OECD average	0.171	0.223	0.291	0.227	0.087	5,201

Table R.8. Students' personal and familiar characteristics. Descriptive statistics. PISA 2000.

Country	Mother works		Father works		Foreign parent ^a	Mother's education		Father's education		100 books ^b
	Women	works	works	works		Secondary	Tertiary	Secondary	Tertiary	
Australia	47.9	71.0	87.1	41.4	41.3	54.5	40.3	54.5	60.5	
Austria	49.4	67.2	92.5	15.0	69.9	22.0	66.2	25.6	45.9	
Belgium	49.4	64.8	89.0	21.4	27.3	65.5	23.9	69.1	40.9	
Canada	49.8	76.9	90.7	20.9	14.7	82.9	21.2	75.2	55.5	
Czech Republic	53.5	83.3	91.9	7.9	40.5	58.0	47.6	50.2	64.4	
Denmark	49.5	84.0	90.4	12.2	32.4	63.1	45.9	48.9	51.6	
Finland	51.6	81.4	86.1	3.2	53.9	40.6	61.9	31.0	44.9	
France	50.6	68.0	88.8	24.2	46.7	45.8	47.5	43.2	42.2	
Germany	50.9	72.2	92.0	20.5	59.3	29.5	47.4	40.5	52.0	
Greece	49.6	47.0	91.0	8.9	45.5	52.6	47.4	50.5	33.0	
Hungary	48.7	72.0	81.2	3.7	42.2	56.7	51.4	46.8	62.1	
Iceland	50.3	86.3	95.6	6.0	52.4	41.2	51.6	42.4	64.4	
Ireland	51.7	61.1	87.9	11.2	38.9	58.8	48.5	48.4	47.0	
Italy	51.0	56.6	93.1	4.8	53.4	43.6	53.2	43.9	46.4	
Korea	44.5	37.8	85.3	0.1	55.1	42.8	47.3	49.8	51.1	
Luxembourg	50.4	54.9	91.5	46.8	47.6	31.3	43.0	37.1	47.5	
Mexico	50.4	38.7	91.5	4.5	64.1	21.2	58.1	29.8	17.6	
Netherlands	49.7	64.3	91.3	15.9	49.1	44.9	39.4	54.0	44.6	
New Zealand	48.9	75.9	90.2	35.9	38.1	46.1	38.0	43.8	54.7	
Norway	49.3	83.1	91.6	11.5	32.0	61.5	34.3	58.3	59.9	
Poland	48.0	61.3	75.7	1.6	34.6	57.3	43.2	47.2	44.0	
Portugal	52.8	68.8	90.9	10.0	68.4	28.7	66.5	29.8	32.3	
Spain	50.4	55.3	91.9	5.6	59.3	33.2	51.9	39.8	51.2	
Sweden	49.1	84.4	90.9	20.0	24.9	70.2	30.4	63.6	61.7	
Switzerland	49.7	64.7	94.8	33.8	69.0	24.8	56.9	37.0	50.3	
United Kingdom	49.5	73.1	87.3	13.4	49.8	43.1	53.1	39.0	46.3	
United States	52.9	77.7	92.2	20.0	13.1	81.6	13.3	79.5	43.7	

Notes: ^a Indicates whether at least one of the parents was born in a country different from their current country of residence or not. ^b Indicates whether there are at least one hundred books at home or not.

Table R.9. Schooling characteristics. Descriptive statistics. PISA 2000.

Country	Funding government ^c	Evaluate teachers ^d
Australia	68.7	27.7
Austria	62.2	18.6
Belgium	67.3	37.6
Canada	34.7	29.8
Czech Republic	74.8	63.6
Denmark	74.6	28.5
Finland	86.1	37.2
France	58.7	37.8
Germany	88.3	11.0
Greece	41.7	21.8
Hungary	57.5	50.5
Iceland	93.6	42.9
Ireland	72.2	38.4
Italy	43.2	77.9
Korea	32.0	24.1
Luxembourg	68.1	38.2
Mexico	58.3	44.5
Netherlands	72.4	40.5
New Zealand	71.4	44.8
Norway	53.5	30.4
Poland	70.9	59.2
Portugal	81.3	29.5
Spain	75.0	41.5
Sweden	41.5	27.5
Switzerland	81.6	8.5
United Kingdom	70.4	61.1
United States	36.9	45.3

Notes: ^c Informs on the percentage of the school total funding for a typical school year that comes from the government. ^d Indicates whether the assessments of students are used to make judgements about teachers' effectiveness or not.

Keywords: Group performance, compulsory education, heterogeneity, categorical data, inverse probability weighting.

JEL codes: I24, C14.

Accepted Manuscript